

# Lyricon:複数アイコンによるビジュアルな音楽選択インタフェース

理学専攻・情報科学コース  
町田和嘉子（指導教員：伊藤貴之）

## 1 概要

近年のインターネット上でのMP3形式音楽配信の急速な普及により、携帯音楽プレイヤーなどで大量の音楽を持ち歩くことが可能になっている。特に最近ではスマートフォンなどの流行により、携帯音楽プレイヤーにおいても選曲や検索のためのユーザインタフェースに関する研究の余地が広がっている。

一方で、リスナーにとって歌謡曲の歌詞は重要な存在である。事前調査として86人の学生に対し「普段歌詞を意識して音楽を聴くことはあるか」というアンケートを実施したところ、66人がYesと答える結果が得られている。よって携帯音楽プレイヤーなどにおいて、歌詞の内容に基づいた選曲を支援するユーザインタフェースがあれば、有用に感じるリスナーは多いであろうと考えられる。

そこで本論文では、楽曲の印象や歌詞の内容に合ったアイコンを、Aメロ、Bメロ、サビなどのブロックごとに選択し、再生時刻順に並べて表示させることで、曲の印象や歌詞のストーリーの時間変化を視覚的に表現するユーザインタフェース“Lyricon”を提案する。この結果を用いれば、実際に曲を聴かなくてもその曲の雰囲気や歌詞内容の展開を想像でき、ユーザは意図通りに楽曲を選択できると考えられる。現在LyriconはWindows PCとAndroidの2種類のプラットフォームに実装されている。

## 2 提案内容

### 2.1 準備処理(1): カテゴリ・キーワードの選定

Lyriconではまず“カテゴリ”として、歌詞のメインテーマになり得る単語を複数用意する。各々のカテゴリは、自身と関連する複数のキーワードと、そのカテゴリを表す複数のアイコンを持つ。本手法ではLyriconにおいてカテゴリとキーワード、アイコンを以下のように定式化する。

- カテゴリを  $C = \{c_1, \dots, c_{N_c}\}$  と定義する。 $N_c$  はカテゴリの総数である。
- $c_i$  に属するキーワードを  $K_i = \{k_{i1}, \dots, k_{iN_{k_i}}\}$  と定義する。 $N_{k_i}$  は  $c_i$  に属するキーワードの総数である。
- $c_i$  に属するアイコンを  $X_i = \{x_{i1}, \dots, x_{iN_{x_i}}\}$  と定義する。 $N_{x_i}$  は  $c_i$  に属するアイコンの総数である。

- アイコン  $x_{ij}$  を表す形容詞を  $A_{ij} = \{a_{ij1}, \dots, a_{ijN_{a_{ij}}}\}$  と定義する。 $N_{a_{ij}}$  は  $x_{ij}$  に割り当てられた形容詞の総数である。

本手法では、86人の学生を対象として、歌詞のテーマに関するアンケートを採取している。その結果をもとに、「恋」「夏」「クリスマス」などの26単語をカテゴリとして採用した。また、無作為選択した歌謡曲100曲の歌詞を形態素解析ソフトウェア茶筌[1]を用いて単語に分割し、日本語ソーラス辞書から抽出したカテゴリの類義語と照合し、共通する248単語を各カテゴリに属するキーワードとして採用した。

### 2.2 準備段階(2): 音楽特徴量の選定

本手法では試験的に無作為抽出した26曲のサンプル曲について、数値解析ソフトウェアMATLAB上に実装された楽曲特徴分析パッケージMIRtoolbox[2]を用いて10種類の特徴量を抽出し、各々に対して2つの相反する意味を持つ形容詞対を定めた。

また本手法では10種類の特徴量の妥当性を検証している。6人の学生被験者に先ほどの26曲のサンプル曲を聴いてもらい、各曲が10種類の形容詞対のそれぞれどちらに該当するかを、被験者らの感性によって選択させた。それと同時に、 $i$ 番目の特徴量  $f_i$  の境界値を  $F_i$  とし、 $f_i$  値が  $F_i$  より大きいかが否かでサンプル曲を2分した。ここで、 $f_i$  に関する2つの形容詞A,Bのうち被験者がA,Bを選択する確率をそれぞれ  $p_A, p_B$  とし、サンプル曲を2分する前のエントロピーを  $E_b(p_A, p_B)$ 、2分した後の各々のクラスタにおけるエントロピーを  $E_{a1}(p_A, p_B)$ 、 $E_{a2}(p_A, p_B)$  とする。本研究では、以下の数式

$$E = E_b(p_A, p_B) - (E_{a1}(p_A, p_B) + E_{a2}(p_A, p_B))$$

で定義される情報利得  $E$  を最大とする  $F_i$  値を求めている。そして、この処理を10種類の特徴量の各々について適用し、情報利得の最大値が高い特徴量3種類を採用した。その特徴量と、割り当てた形容詞を表1に示す。

表 1: 採用した特徴と形容詞

特徴	形容詞
テンポ (tempo)	遅い・速い
高音域の割合 (brightness)	単純な・輝かしい
不調和音の多さ (roughness)	素朴な・複雑な

### 2.3 楽曲入力時の処理 (1): アイコン候補の選定・列挙

本手法では各カテゴリに対し、前節で選定した形容詞に対応する 6 種類のアイコンを選定した。アイコンは、12 人の学生被験者に「アイコン画像から受けるイメージ」を回答させ、我々の意図するイメージと合致した結果を得たものを採用した。そして各アイコンには、前節で選定した形容詞をタグとして付与する。

アイコンを曲のブロックごとに選択するに先立ち、Lyricon は最初に歌詞をブロックごとに分割する。本手法では歌詞のブロック分割に LyricMaster[3] を用いている。続いてブロックごとに分割した歌詞に茶筌 [1] を用いて形態素解析を適用し、単語の最小単位に分割する。ここで、ある 1 つのブロックに含まれていた単語を  $W = \{w_1, \dots, w_N\}$  とする。もし単語  $w_k$  がキーワード  $k_{ij}$  と一致したら、Lyricon はこのブロックをカテゴリ  $c_i$  と関連付ける。この場合、Lyricon はアイコン群  $X_i$  に含まれる複数のアイコンを表示用アイコンの候補とする。

### 2.4 楽曲入力時の処理 (2): アイコンの最終決定

Lyricon は続いて、2.2 節で選定した 3 つの特徴量を算出し、以下の手順でその曲への相応しさの順位を決定する。まず、ここで扱う数値を以下のように定式化する。

- 1.2 節で求めた境界値を  $F_i$  とする。  
( $i = \{tempo, brightness, roughness\}$ )
- 1.2 節で求めたサンプル曲の各特徴量の最大値と最小値の差を  $D_i$  とする。
- 算出した音楽特徴量を  $M_i$  とする。

以上の数値から Lyricon では、特徴の突出度  $R_i = \frac{|M_i - F_i|}{D_i}$  を求める。そして、 $R_i$  の値を最大にする  $M_i$  に対応する形容詞と同じ形容詞が設定されているアイコンを各ブロックにおける最終選択結果とする。

## 3 ユーザーインターフェース

本手法では、Lyricon によって自動選択されたアイコンを一覧表示するプレイリストを、Windows PC と Android

携帯エミュレータ上で実装している。現時点での実装では、アイコン画像には、横 32 × 縦 26 ピクセルのものを用いている。なお、両実装とも曲の再生・停止のみならず、曲の頭出しや途中再生も可能である。

まず Windows PC での実装画面を図 1 に示す。この実装画面では、縦方向に楽曲群が、横方向に 1 曲のアイコン群が並んでいる。

続いて、Android 携帯エミュレータでのユーザーインターフェースの表示例を図 2 に示す。この実装では、縦方向のスクロールバーによって多数の曲を眺め、横方向のスクロールバーで曲全体を通したアイコン群を眺めることができる。また、タイトルボタンを押すことにより、図 2(右) に示すように、その曲全体を表現するアイコンを表示する。

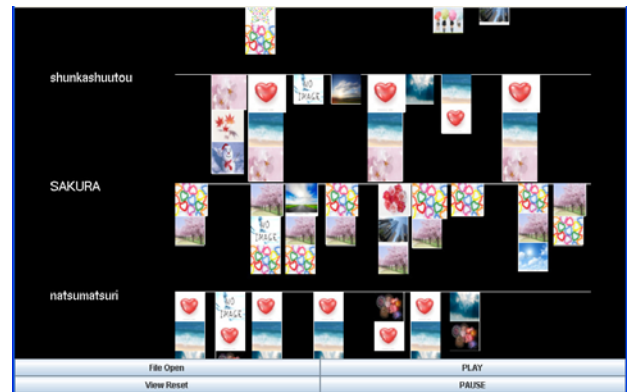


図 1: Windows PC での実装画面

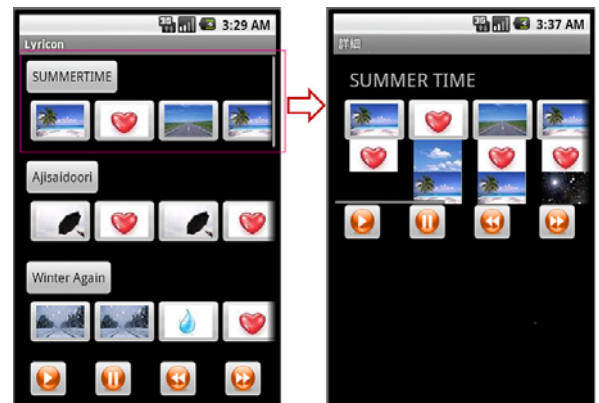


図 2: Android 携帯エミュレータでの実装画面

## 参考文献

- [1] 茶筌, <http://chasen.naist.jp/hiki/ChaSen/>
- [2] MIRtoolbox, <https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox>
- [3] Lyric Master, <http://www.kenichimaehashi.com/lyricsmaster/>